# INDUSTRIAL PROJECTS 234313

## WINTER 2016-17

<table>
<thead>
<tr>
<th>No.</th>
<th>Project Domain</th>
<th>Project Name</th>
<th>Company</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Video player extensions</td>
<td>Player Plugins</td>
<td>AOL</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>Embedded real time operating systems, HLS.</td>
<td>RTOS SOFT-CORE Accelerator</td>
<td>Elbit Systems Ltd</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>Medical imaging, networks, streaming services</td>
<td>Rapid medical information communication over slow networks</td>
<td>Philips Medical Systems</td>
<td>7</td>
</tr>
<tr>
<td>4</td>
<td>SW systems, informatics, analytics</td>
<td>Intelligent documents proofing and indexing system</td>
<td>Philips Medical Systems</td>
<td>8</td>
</tr>
<tr>
<td>5</td>
<td>Enterprise systems, machine learning, analytics</td>
<td>A machine learning based system for automatic service tickets classification</td>
<td>Philips Medical Systems</td>
<td>9</td>
</tr>
<tr>
<td>6</td>
<td>NLP, machine learning</td>
<td>Textual data Representation for Medical Images classification</td>
<td>Philips Medical Systems</td>
<td>10</td>
</tr>
<tr>
<td>7</td>
<td>D.B / web</td>
<td>Html5 registers functional specification</td>
<td>Marvell</td>
<td>11</td>
</tr>
<tr>
<td>8</td>
<td>Computer architecture optimization</td>
<td>Consumer producer workload optimization</td>
<td>Intel</td>
<td>11</td>
</tr>
<tr>
<td>9</td>
<td>Image processing, physics engines, VR</td>
<td>Physics – Based Virtual Camera</td>
<td>Intel</td>
<td>12</td>
</tr>
<tr>
<td>10</td>
<td>System programming</td>
<td>Application-specific data-structure visualization framework for Windows debuggers</td>
<td>Mellanox</td>
<td>13</td>
</tr>
<tr>
<td>11</td>
<td>Resiliency testing</td>
<td>Chaos Events Simulation</td>
<td>HPE Software</td>
<td>15</td>
</tr>
<tr>
<td>12</td>
<td>Systems</td>
<td>Selenium Run Engine for AppPulse</td>
<td>HPE Software</td>
<td>16</td>
</tr>
<tr>
<td>Column 1</td>
<td>Column 2</td>
<td>Column 3</td>
<td>Column 4</td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>---------------------------------------------------------------------------</td>
<td>------------------</td>
<td>----------</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Mobile, web development</td>
<td><strong>Active</strong></td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Java systems</td>
<td><strong>BuiltWith for Android</strong></td>
<td>HPE Software</td>
<td>17</td>
</tr>
<tr>
<td>15</td>
<td>Image analysis, augmented reality, mobile dev</td>
<td><strong>AppPulse Chaos Agent</strong></td>
<td>HPE Software</td>
<td>17</td>
</tr>
<tr>
<td>16</td>
<td>Data mining</td>
<td><strong>Smart Editor for Home Wall Design</strong></td>
<td>Shutterfly</td>
<td>18</td>
</tr>
<tr>
<td>17</td>
<td>User interaction simulation, data mining</td>
<td><strong>Face recognition evaluation framework</strong></td>
<td>Shutterfly</td>
<td>19</td>
</tr>
<tr>
<td>18</td>
<td>Data mining</td>
<td><strong>Testing framework simulating user account</strong></td>
<td>Shutterfly</td>
<td>20</td>
</tr>
<tr>
<td>19</td>
<td>Data collection</td>
<td><strong>User account analysis</strong></td>
<td>Shutterfly</td>
<td>21</td>
</tr>
<tr>
<td>20</td>
<td>Face recognition</td>
<td><strong>Fast mobile share</strong></td>
<td>Shutterfly</td>
<td>22</td>
</tr>
<tr>
<td>21</td>
<td>Mobile/web programming</td>
<td><strong>Social media to photo product</strong></td>
<td>Shutterfly</td>
<td>23</td>
</tr>
<tr>
<td>22</td>
<td>Machine learning</td>
<td><strong>Smart notification server</strong></td>
<td>Shutterfly</td>
<td>23</td>
</tr>
<tr>
<td>23</td>
<td>Data collection</td>
<td><strong>Shipment tracking agent</strong></td>
<td>Shutterfly</td>
<td>24</td>
</tr>
<tr>
<td>24</td>
<td>stream processing</td>
<td><strong>Real-time aggregation of user interaction data using Kafka Streams</strong></td>
<td>Outbrain</td>
<td>25</td>
</tr>
<tr>
<td>25</td>
<td>Code generation, computer languages</td>
<td><strong>UI Coder</strong></td>
<td>Rafael</td>
<td>26</td>
</tr>
<tr>
<td>26</td>
<td>Management</td>
<td><strong>google test launcher</strong></td>
<td>Rafael</td>
<td>27</td>
</tr>
<tr>
<td>27</td>
<td>Communication</td>
<td><strong>XDP 2.0 Communication Protocol</strong></td>
<td>Rafael</td>
<td>27</td>
</tr>
<tr>
<td>28</td>
<td>Image generation and Image processing</td>
<td><strong>Augmented Reality</strong></td>
<td>Rafael</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>Topic</td>
<td>Description</td>
<td>Provider</td>
<td>Page</td>
</tr>
<tr>
<td>---</td>
<td>-----------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------</td>
<td>----------</td>
<td>------</td>
</tr>
<tr>
<td>29</td>
<td>Databases, parallel and distributed programming</td>
<td>SiFra-DisCoM (Simulation Framework for Distributed Court Manager)</td>
<td>IBM</td>
<td>29</td>
</tr>
<tr>
<td>30</td>
<td>Document classification</td>
<td>Document Classification for regulatory compliance</td>
<td>IBM</td>
<td>31</td>
</tr>
<tr>
<td>31</td>
<td>Adaptive bitrate video streaming, HLS, DASH</td>
<td>Live video crawler using the HLS and DASH adaptive bitrate video streaming standards</td>
<td>IBM</td>
<td>32</td>
</tr>
<tr>
<td>32</td>
<td>Cloud computing, big data</td>
<td>Serverless Big Data Analytics</td>
<td>IBM</td>
<td>33</td>
</tr>
<tr>
<td>33</td>
<td>Cloud systems, security, machine learning</td>
<td>Cognitive Cloud Shield</td>
<td>IBM</td>
<td>34</td>
</tr>
<tr>
<td>34</td>
<td>Cloud, PAAS, Docker</td>
<td>Video surveillance service using FFMPEG and Docker on IBM Bluemix cloud</td>
<td>IBM</td>
<td>35</td>
</tr>
<tr>
<td>35</td>
<td>Open source, WebRTC, Unified communication, IoT</td>
<td>Bridge the gap between different communication hubs</td>
<td>Amdocs</td>
<td>37</td>
</tr>
<tr>
<td>36</td>
<td>Open Source, WebRTC, unified communication, IoT</td>
<td>Bridge the gap between different management tools</td>
<td>Amdocs</td>
<td>38</td>
</tr>
<tr>
<td>37</td>
<td>Open Source, WebRTC, Unified communication, IoT</td>
<td>Where do you store your images? Why should it matter! Bridge the gap between different media services</td>
<td>Amdocs</td>
<td>40</td>
</tr>
<tr>
<td>38</td>
<td>Open Source, WebRTC, Unified communication, IoT</td>
<td>Connect Facebook Messenger and WhatsApp: Bridge the gap between different messaging services</td>
<td>Amdocs</td>
<td>42</td>
</tr>
<tr>
<td>39</td>
<td>Big Data, analytic, diagnostic, REST</td>
<td>Big Data - Diagnostic package collector</td>
<td>Amdocs</td>
<td>44</td>
</tr>
<tr>
<td>40</td>
<td>Predictive analytics algorithms,</td>
<td>From Information to Optimization by</td>
<td>Amdocs</td>
<td>45</td>
</tr>
<tr>
<td>Big Data</td>
<td>Predictive analytics algorithms</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------------------------------------</td>
<td>---------------------------------</td>
<td>------------------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>41 Canceled</td>
<td>Canceled</td>
<td>Amdocs</td>
<td>47</td>
<td></td>
</tr>
<tr>
<td>42 Integration hub with multiple chatbots providers</td>
<td>Chatbots Hub</td>
<td>Amdocs</td>
<td>47</td>
<td></td>
</tr>
<tr>
<td>43 web application development mechanism</td>
<td>Text Provider for web applications</td>
<td>Amdocs</td>
<td>48</td>
<td></td>
</tr>
<tr>
<td>44 Web and mobile UI, JavaScript application.</td>
<td>Fluid Page Layout without WCMS – Dynamic UI presentation</td>
<td>Amdocs</td>
<td>49</td>
<td></td>
</tr>
<tr>
<td>45 limited time promotions using Snapchat</td>
<td>Snapchat advertisement platform</td>
<td>Amdocs</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>46 Web/Mobile &amp; analytics application</td>
<td>SMB Marketplace Employee Cockpit</td>
<td>Amdocs</td>
<td>51</td>
<td></td>
</tr>
<tr>
<td>47 Cognitive analytics, machine learning, automation, AI, Robotic Process Automation</td>
<td>Cognitive Automation for Productivity (CAP system)</td>
<td>Amdocs</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>48 Machine Learning, Databases</td>
<td>Declarative System for Engineering Deep Learning</td>
<td>LogicBlox</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td>49 Machine Learning, Data Mining</td>
<td>Tool for Data Quality Management in Machine Learning</td>
<td>LogicBlox</td>
<td>57</td>
<td></td>
</tr>
</tbody>
</table>
PROJECT 1

Company: AOL

Project Domain: Video player extensions

Project: Player Plugins

Supervisor: Ady Levy

Project Description: Ability to develop any add-on with / without UI on top of the AOL’s player which can reach millions of users every day.

Some examples:
- Recommendation engine, with UI
- Gather Quality analytics
- Voting and polls add-ons
- Cutting edge ad formats

Programming Languages and Development Platform: Vanila Javascript, Web and development tools

Course: 234319

PROJECT 2

Company: Elbit Systems

Project Domain: Embedded Real Time Operating Systems, HLS.

Project: RTOS SOFT-CORE Accelerator

Supervisor: Eran Peled
Background: A real-time operating system (RTOS) is an operating system (OS) intended to serve real-time application process data as it comes in, typically without buffering delays. Latency and determinism are two of the most important aspects in designing a RTOS. Today’s technology and tools enables the creation of HW accelerator IP using HLS – High level Synthesis. C/C++ software language can be synthesized into Programmable devices without the need to manually create RTL.

Using this technology enables us to create a hybrid hardware/software RTOS that takes advantage of hardware accelerators to improve the critical metrics and, consequently, improve overall system performance.

Project Description: In this project, the student will work with a development board Xilinx ZC702 which includes a XILINX ZYNQ7020 SOC (System on Chip). This SOC includes 2 ARM Cortex-A9 and an FPGA.

The project will comprise:
1. Learning basic operations and flow of the XILINX tools: VIVADO, VIVADO HLS, and VIVADO SDK.
2. Use the tools to create and run (On the ARM core) the built-in hello world examples supplied by Xilinx.
3. Offer and implement (in SW) a very primitive scheduler which includes cyclic/upon-event tasks without preemption (Preemption raises the complexity by far).
4. Implement this scheduler or selected parts of it as an HW accelerator using HLS tools.
5. Create a report summarizing the comparison between the HW and SW scheduler the consideration for the HW/SW logic split and a How-To guide of the OS creation design.
6. Optional as a dual-semester project: perform the same operations with FreeRTOS (open source RTOS).

Programming Languages and Development Platforms: C, HLS, VHDL (advantage).
Courses:
- 234114/7  Introduction to Computer Science
- 234122  Introduction to Systems Programming
- 234123/046209  Operating Systems
- 234267/046267  Digital Computers Structure (MAMAS)

PROJECT 3

Company: Philips Medical Systems
Project Domain: Medical imaging, networks, streaming services
Project: Rapid medical information communication over slow networks
Supervisor: Moshe Katz

Project Description: Medical imaging is a domain involves processing of huge amount of information under tight performance constraints and is limited today by network conditions. In order to work with a community of professionals distributed geographically or to provide medical solutions as a service, it is crucial to provide a fast, reliable and efficient means to distribute the data.

The purpose of this project is to develop a medical communication channel on top of advanced streaming and files transfer services usually used for multi-media content transfer.

The project will involve learning of the relevant protocol basics, network principles, evaluation of various streaming services and the development of endpoints on top of the selected streaming service API.

Programming Languages and Development Platform: C# as the basic programing language, medical communication toolkits, other programing language and platforms according to the selected streaming service.
PROJECT 4

Company: Philips Medical Systems

Project Domain: SW systems, informatics, analytics.

Project: Intelligent documents proofing and indexing

Supervisor: Moshe Katz

Project Description: In many industries, the development of a new product involved authoring and archiving of a lot of requirements, design and planning documents. These documents follow formal standards and templates which has to be strictly followed and are being stored in archives that are designed to maintain projects’ structure, version management and traceability but much less for contextual informatics and content retrieval.

The purpose in this project is two-fold:

- Automatic proofing of documents to make sure they are compliance with the standard according to various templates and pre-defined rules in order to save a lot of erroneous manual reviews
- Keywords extraction and indexing of the documents according to dynamic weights of keywords derived from documents structure

The project is flexible in terms of the programming language and will involve different techniques, analytics tools and frameworks.

Programming Languages and Development Platform: Flexible selection of programing languages, Elastic Search for indexing.
PROJECT 5

Company: Philips Medical Systems

Project Domain: Enterprise systems, machine learning, analytics

Project: A machine learning based system for automatic service tickets classification

Supervisor: Moshe Katz

Project Description: At the center of any enterprise product value chain stands the quality of the ongoing service. This is also the source of a large portion of product’s delivery cost. This project is aimed to provide an automatic analysis and classification of service calls retrieved from the CRM system in order to identify actual importance and shorten the time to resolution.

The idea is to correlate between specific call characteristics, previous similar calls and relevant defects (bugs) tracked in the defects management system. Once identifying similar issues, important associated with them, indications of whether they have been addressed and what kind of solution has been given will optimize the capacity allocated to resolve the problem.

The project involves basic integration with management systems, usage of natural language processing algorithms, machine learning and various analytics techniques, using Apache Sparks and Elastic Search.

Programming Languages and Development Platform: C# or Java, Apache Spark and Elastic Search
Company: Philips Medical Systems
Project Domain: NLP, Machine Learning
Project: **Textual data Representation for Medical Images classification**
Supervisor: Bella Fadida-Specktor

**Project Description:** Classification of medical images based on DICOM (Digital Imaging and Communications in Medicine) textual metadata can serve multiple purposes. In this work we will focus on predicting which preprocessing algorithm needs to be applied on a given data type. The aim of this project is to develop a textual representation of DICOM metadata which will then be used for data classification.

The project includes building and testing several textual representations and using them as an input to an existing machine learning framework for Pre-processing prediction. The final goal will be comparing classification accuracy using different representation types.

**Programming Languages and Development Platform:** Java, Apache Spark MLlib library

**Recommended Background:** NLP and/or Artificial Intelligence group of courses.
PROJECT 7

Company: Marvell
Project Domain: D.B / WEB
Project: Html5 registers functional specification
Supervisors: Idit Schwarts and Yehuda Ganel

The project's goal is to create a tool whose purpose is to convert register information from SOC's XML format to the HTML format. This will enable a more informative and intelligent presentation of the registers, using complex GUI objects such as navigational trees, tables, frames, etc. The tool will serve as a modern alternative to the current methods of presentation (PDF documents).

Programming Languages and Development Platform: JS, XSLT, html, XML, C#
Courses: Introduction to computer science, software engineering

PROJECT 8

Company: Intel
Project Domain: Computer architecture optimization.
Project: Consumer producer workload optimization
Supervisors: Yoni Aizik and Efi Rotem

Project Description: In modern microprocessor, often the performance of an application is limited by the power that the CPU can dissipate. On a specific type of application, consumer-producer, two threads, cores, or components such as graphics processor are working together...
on a single task. Optimizing the overall performance of such a task under system constraints is a challenge. The student is required to write a framework that generates and runs synthetic consumer-producer workloads, and will optimize the performance on an Intel Skylake system using micro architectural parameters.

The project includes:

1. Writing a framework to generate consumer-producer workload between single core and many cores and between single core and gfx using OpenCL that can:
   - control the duration of each portion of the work
   - control the memory access characteristics of the IA workload
   - control the number of execution units running simultaneously (AVX 256, integer, number of execution units in GFX)

2. Optimizing these synthetic workloads on a Skylake system.

**Programming Languages and Development Platform:** Main work: C/C++; Optional: OpenCL, Python

---

**PROJECT 9**

**Company:** Intel

**Project Domain:** Image Processing, Physics Engines, VR

**Project:** Physics – Based Virtual Camera

**Supervisor:** Gilad Bauman

**Project Description:** Many virtual reality and computer graphics applications require the use of one or more virtual cameras which define the point of view from which we observe the virtual world. There are many methods for controlling that camera: from keyboards, mice and gamepads to full head tracking in virtual reality. What these methods have in common is that
they attempt to bridge the gap between the human and the virtual world. However, due to the complexity of human perception and the way humans interact with the real world, it is difficult to fully model camera movement in a virtual environment in a way that is both intuitive and pleasing to experience. Our goal is to enable exactly such an experience for viewing 3D virtual worlds by applying physical qualities to the virtual camera. This will allow us to express our requirements as natural movements with the Newtonian physics laws of motion. In this project we will model a physics-based virtual camera which will be controlled by dynamic forces and kinematic constraints.

**Programming Languages and Development Platform:** C/C++/C#/...

**Course:** 236816

---

**PROJECT 10**

**Company:** Mellanox

**Project Domain:** System programming

**Project:** Application-specific data-structure visualization framework for Windows debuggers

**Supervisor:** Oved Itzhak

**Project Description:** The traditional approach for application-specific data structure visualization is to develop application-specific debugger extensions.

Consider a dictionary data-structure that is implemented as a Binary-Search-Tree (BST) – inspecting the dictionary by a programmer through a debugger is best achieved by showing it as a dictionary – a linear list of keys and values, ordered by the same ordering criteria as the tree. However, in general a debugger cannot automatically deduce that a certain data structure is a node in a binary tree (let alone in a BST). At most it can allow manual step through pointers. Moreover, even if the debugger could employ some heuristic (e.g. if a structure contains 2
pointers to the type of itself which is quite contrived…) it cannot automatically deduce the key, let alone how to display it.

This is where a debugger extension comes in. It is developed as part of the application, specifically for the purpose of debugging. It is where the application-specific knowledge is coded.

The Windows debuggers provide integration interface and services for such extensions. These extensions are packaged as Windows DLLs. They are invoked by the debugger and live in the debugger program’s address-space, which is separate from the debuggee application’s address-space. This makes going over the a data structure in the debuggee completely different from going over the data-structure in the same address-space – the extension needs to copy memory blocks from the debuggee address-space to the debugger’s and parse them. For example, pointers obtained from the debuggee must not be dereferenced but rather to inspect the pointed data the respective memory in the debuggee need to be copied to the debugger.

The Microsoft Visual C++ compiler contains non-standard language extension that allows encapsulating this functionality in an in-process like interface, i.e. dereferencing a debuggee pointer in the extension implicitly initiate copying and parsing the pointed data in the debuggee. This greatly simplifies developing visualizers.

In this project we will implement a framework, consisting of a code generator and runtime support, which automates the implementation of this technique, leaving only the data rendering logic to the debugger extension author to implement.

**Programming Languages and Development Platform:** Language: C++; Platform: Windows

**Courses:** 234122 - Introduction to Systems Programming; 234123 - Operating Systems
Company: HPE
Project Domain: Resiliency Testing
Project: Chaos Events Simulation
Supervisor: Avivi Siman-Tov

Project Description: One of the big challenges when testing a system for resiliency is to be able to simulate a failure rather than creating it for real. This way allows to maintain the system functional and available for other purposes.

The goal of this project is to suggest a solution / approach for simulating different kind of failures at different levels starting from a hardware component that fails (e.g. HDD) and ending with a specific service.

Programming Languages and Development Platform: Java

Courses:
- Programming languages and development platforms
- Introduction to Computer Science
- Computer Organization and Programming
- Introduction to Systems Programming
- Operating Systems
- Data Structures 1
- Algorithms 1
- Programming Languages
- Information Storage Systems
- Introduction to Computer Networks
- Computer Security
- Database Management Systems
- Operating systems engineering
- Software design
- Object Oriented Programming

PROJECT 12

Company: HPE Software
Project Domain: Systems
Project: Selenium Run Engine for AppPulse Active
Supervisor: Aviad Israeli and Amichai Nitsan

Project Description: AppPulse Active is a product that is actively checking the health of web applications and reports the app owners when there is a critical failure with their apps. It is doing so by running certain routine checks, written as scripts, and checks the results of these tests.

In this project the students will develop an engine on a Linux environment that is able to run scripts developed with “Selenium”, a popular framework for web automation. This engine will be able to receive directions from remote and be responsible for scheduling the runs in the proper time and also collect the results and send them back. The engine will use one of the cutting edge technologies for software containerization – “Docker containers” to run each script in an isolated sandbox.

Programming Languages and Development Platform: Good knowledge in Java and JavaScript is required.

PROJECT 13
**Company**: HPE Software  
**Project Domain**: Mobile, Web Development  
**Project**: BuiltWith for Android  
**Supervisors**: Avi Kabizon and Amichai Nitsan  
**Project Description**: “BuiltWith” is a well-known tool that analyses a web site and tells you what technology components are used to build it.

This project is about building a similar tool to BuiltWith, but for android developers. The tool will be able to take an APK (android application file) and report what components, or 3rd party libraries, are present in this APK.

The second part of the project is to build a website where users can upload APKs or choose applications from the Google Play store.

An example for a similar service can be found here: [https://www.apptopia.com/](https://www.apptopia.com/)  
And the result page: [https://www.apptopia.com/apps/google_play/com.tranzmate/sdks](https://www.apptopia.com/apps/google_play/com.tranzmate/sdks)

**Programming Languages and Development Platform**: The APK analysis part will be developed in Java. The website should be developed in any modern web development environment – to be chosen together by the students and instructor.

---

**PROJECT 14**

**Company**: HPE Software  
**Project Domain**: Java Systems  
**Project**: AppPulse Chaos Agent  
**Supervisors**: Merav Aizenfeld and Amichai Nitsan  
**Project Description**: AppPulse Trace is a tool for monitoring systems in production, identifying problems and offering tools to developers for quickly fixing the problems. Our agents are
installed on customer servers and everything can happen there. We need to do our best for being ready to every possible server environment.

In this project, the students will develop the “Chaos Agent”. It is a Java library that can be included with any Java application. The purpose of this agent is to generate chaos, to generate problematic situation of anything bad that can go on our customer environment in which Java application it is hosted inside. From issues specific to the application like a lot of threads, through slow networks and lack of memory.

Using the chaos agent, the Java applications will have simulated problems that our tool could detect and alert on.

Such failures could be slow network, overused memory or unexpected exceptions.

**Programming Languages and Development Platform:** Good knowledge of Java is needed, and some understanding in communication, databases and operating systems can help.

---

**PROJECT 15**

**Company:** Shutterfly

**Project Domain:** Image Analysis, Augmented reality, Mobile dev.

**Project:** Smart Editor for Home Wall Design

**Supervisor:** Roman Sandler

**Project Description:** Shutterfly suggests its customers a wide range of types of photo products. One of them is wall décor products. A customer may design a real wall decoration for his/her house, by ordering a set of canvas or metal prints that will fit his/her taste. This process however, is very time consuming and requires a lot of patience.
It turns out that our favorite photos, even taken by 20 megapixel cameras, are not always good enough to be printed in high resolution. Moreover, after choosing one good photo it is hard to find 2-3 more in your 50000+ photos that will fit it well on the wall of your dining room.

The initial step of the project is shouting the wall the user wishes to decorate. Then a wall design based on analysis of thousands of user’s photos is suggested. The design will be presented in simple mobile application. The next step will be to create an assistant that after the change one of the photos will provide several suggestions for design update.

**Programming Languages and Development Platform:** Mobile development (OS by preference of the students)

**Courses:** Course in image analysis is an advantage

### PROJECT 16

**Company:** Shutterfly

**Project Domain:** Data mining

**Project:** Face recognition evaluation framework

**Supervisor:** Omer Geiger

**Project Description:** Shutterfly provide its users with tools to simplify photo management experience. One of them is face recognition for all user photos. A customer may store hundreds of thousands of his photos in Shutterfly account. One way to simplify finding the right photo is automatic face recognition across the account. Developing such ability requires a framework which is a challenge by itself.

The outcome of this project is a framework that allows to simulate face recognition process on a user's account and to compare different scenarios and algorithms. Using this framework, several
experiments will be conducted by the students to compare the performance of different algorithms and parameter configurations in terms of precision and recall

**Programming Languages and Development Platform:**
- Web technology, Php/python, AWS.
- Face recognition algorithms

**Courses:** Course in AI is an advantage

---

**PROJECT 17**

**Company:** Shutterfly

**Project Domain:** User interaction simulation, data mining

**Project:** Testing framework simulating user account

**Supervisor:** Itzik Shabtay

**Project Description:** Shutterfly suggests its customers a wide range of types of personalized photo products. Testing those products is a nontrivial task since most of the flow is user specific. One of the hardest challenges in testing customer experience is recreation the very specific concurrence of customer steps and backend processes that cause wrong system behavior. The goal of this project is to create a framework that allows simulating a large variety of user behaviors and recreating the sequence of events that caused an error. Specifically, it will create a wide range of photo collections from a photo set, create user accounts varying user answers during account creation, simulate different scenarios of photo upload, and simulate different user behavior as a response to events on the site. The goal of the tool will be also to narrow down the parameters that result in undesirable program behavior (not necessarily a crash).

**Programming Languages and Development Platform:** Web development.

**Courses:** Course in AI is an advantage
PROJECT 18

**Company:** Shutterfly  
**Project Domain:** Data mining  
**Project:** User account analysis  
**Supervisor:** Roman Sandler  
**Project Description:** No customer is eager to share his information with service provider. It is an annoying process with no immediate positive impact on the customer experience. It is even more annoying when the user perceives the requested information as the one already available to the service provider.

The goal of this project is to create a user management console that unifies both explicit and implicit information available from different data sources. The missing information will be predicted using tools of machine learning.

**Programming Languages and Development Platform:** Any method convenient to the students to proceed and summarize large amounts of data

PROJECT 19

**Company:** Shutterfly  
**Project Domain:** Data collection  
**Project:** Smart collection of mobile behavior info for experience personalization  
**Supervisor:** Roman Sandler  
**Project Description:** Every customer browses the site/application of retailer before purchasing. Some retailers use the history of such browsing to push additional products that seem to interest the customer. In physical world some retailers even make the customer pass those
products before finding the product the customer is really looking for, hoping for impulse purchase.

In this project we intend to collect and A/B test browsing paths of users in mobile application. The goal is to summarize the browsing info into short, comparable string of items, train machine learning algorithm on it and optimize user experience. The resulting browsing path should be shorter than today, but let the user to consider all purchases he wants.

**Programming Languages and Development Platform:**

- Any method convenient to the students to proceed and summarize large amounts of data.
- Mobile development

---

**PROJECT 20**

**Company:** Shutterfly

**Project Domain:** Face recognition

**Project:** **Fast mobile share**

**Supervisors:** Omer Geiger and Yoni Mor

**Project Description:** Shutterfly provide its users with tools to simplify photo management experience. One of such tools is face recognition. Many of the contacts in today's mobile phones are associated with person's photo.

The task of this project is to create an application that analyzes photos taken with it and proposes to share the good photos with the contacts appearing in the image.

**Programming Languages and Development Platform:** Mobile development.
PROJECT 21

Company: Shutterfly

Project Domain: Mobile/Web programming

Project: Social media to photo product

Supervisors: Yohan Sabbah and Yoni Mor

Project Description: Social media is very convenient for fast sharing of photos with friends and family, but not for creating the photo products (e.g. photobooks, greeting cards, etc.) from them. The task of this project is to connect to user's accounts in SM and suggest Shutterfly photo products based on images shared in specific event or a series of events. The main aspects of the projects are programming using SM and realizing the context of the share to choose the relevant products

Programming Languages and Development Platform: Mobile or web development

PROJECT 22

Company: Shutterfly

Project Domain: Machine learning

Project: Smart notification server

Supervisor: Yoni Mor

Project Description: No customer likes to be over-notified. Endless emails, pop ups, notification balloons annoy people. No customer likes to be under-notified. People opt out from e-mail subscriptions and then call customer support to complain that they were not notified.

In this project we shall try to build a framework for fine scale tuning for the notifications we wish to show to our customers and the situations in which we should really do so.

Programming Languages and Development Platform:
- Any method convenient to the students to proceed and summarize large amounts of data.
- Mobile development.

**PROJECT 23**

**Company:** Shutterfly

**Project Domain:** Data collection

**Project:** Shipment tracking agent

**Supervisor:** Yoni Mor

**Project Description:** Shutterfly sends every year hundreds of millions of parcel deliveries. The operation includes Shutterfly employees who should deliver the mail, delivery service providers, call center employees, and, of course, the customers. In principle the delivery providers provide tracking interface for every order number and the raw information is available to the customer. However, we are all people, and still curious for more data and have strange requests. The goal of this project is to create an agent that oversees the delivery, compares it to the statistics of similar deliveries, looks for other deliveries of the same user that could be merged, and if needed keeps the customer informed about these details.

**Programming Languages and Development Platform:** Web development
PROJECT 24

**Company:** Outbrain

**Project Domain:** stream processing

**Project:** Real-time aggregation of user interaction data using Kafka Streams

**Supervisor:** Ronny Lempel

**Project Description:** The project builds a very common capability of large-scale Web companies using some of the latest big-data technologies. Specifically, the project will push (simulated) streams of user impressions and clicks on many different documents into a Kafka queue. Inside Kafka, the newly released Kafka Streams library will be used to aggregate separately impressions and clicks, per document, per 2-minute time windows. Those values will be written to Cassandra, a scalable NoSQL database that will be used here as a key-value store. A separate service will then query Cassandra, given a document identifier and a timeframe, and will plot a graph of the click-through rate of the document over the timeframe.

**Programming Languages and Development Platform:**
- The project will be developed in Java or Scala.
- Will require usage of Kafka 0.10: [http://kafka.apache.org/downloads.html](http://kafka.apache.org/downloads.html)

**Courses:**
- 236370 Concurrent and Distributed Programming
- 236363 Database Management Systems
- 236620 Big Data Technologies
PROJECT 25

**Company:** Rafael

**Project Domain:** Code Generation, Computer Languages

**Project:** UI Coder

**Supervisor:** Yariv Kimchi

**Project Description:** In our simulation platform (DSIDE) we built a descriptive API to build and design systems and models:

```csharp
void CreateSystem()
{
    var tank = PlaceModel<Tank>.InRoot();
    var pedestal = PlaceModel<Pedestal>.On(tank);
}
```

After the system is built and the relationship of a system is defined the models can then be connected using events:

```csharp
void InitializeSystem()
{
    var tank = FindModel<Tank>();
    var pedestal = FindModel<Pedestal>();

    tank.OnChangedPosition += SetPedestalOrientation(tank.Position);
}
```

In a nutshell - the project goal is to build a UI which will allow us to describe a declarative language (like UML or simulink) that will generate code to look like the above example – the
project will be a base to a broader family of code generation tools, and should be adaptable to any API.

**Programming Languages and Development Platform:** Visual Studio, C#, WPF (or any other).

---

**PROJECT 26**

**Company:** Rafael  
**Project Domain:** Management  
**Project:** Google test launcher  
**Supervisor:** Yehuda Aharon Ezer

The project manager: The manager oversees the execution of unit tests written in gtest, and is responsible for generating and printing the list of existing tests, enabling and ordering unit test executions, and setting dependencies between tests.

The manager is responsible for remote execution of tests, meaning to send the test files to a remote computer, run them there, and send the results to the local computer to display them to the user.

The manager must ensure that test execution time does not exceed the defined time for each test and display a report of successfully run tests / tests that failed and generate a word document.

**Programming Languages and Development Platform:** Java over windows

---

**PROJECT 27**

**Company:** Rafael  
**Project Domain:** Communication  
**Project:** XDP 2.0 Communication Protocol
Supervisor: Yehuda Aharon Ezer

The project aims to define and implement retransmit functionality, and to demonstrate a system that simulates hypothetical events and tests the protocol's interaction with those events.

The protocol is designed for UDP without breaking the communication stream and real-time for CPU-intensive applications that require full control over the transferred packets.

Programming Languages and Development Platform: C++ in Visual Studio

---

Company: Rafael

Project Domain: Image Generation and Image Processing

Project: Augmented Reality

Supervisor: Diana Pekerman

The project goal is to combine virtual objects with live video when given camera data. The process involves creating a virtual image on the fly and adjust the positions and movements of the virtual objects.

Programming Languages and Development Platform: PC; Languages – C#, C++, Matlab

Courses:
- 3D Graphics
- Computer Animation

---

Page 28 of 59
PROJECT 29

Company: IBM

Project Domain: Databases, Parallel and Distributed Programming

Project: SiFra-DisCoM (Simulation Framework for Distributed Court Manager)

Supervisor: Yan Tzitrin

Project Description:

The need for court case management software is widely discussed in the literature and the world wide web (see, http://www.capterra.com/court-management-software/ for example). Various approaches were proposed in order to provide the customers with a comprehensive, high-quality, efficient and convenient tools (e.g. https://www.court.gov.il/, e-Court etc.).

We are developing DisCoM (Distributed Court Manager) - a platform for a distributed court management. The basic description and requirements for this platform are summarized in Figure 1).
Figure 1: (a) **Assumptions**: [A1] The DisCoM clients (Parties, for example Low Offices) in general case reside on different sites; [A2] The parties send request to the DisCoM regarding their juridical case (JC): create, update, delete; [A3] The parties’ times are not synchronized; [A4] In general case, the requests related to a juridical case (even from the same Party) are not delivered to DisCoM in their sending order.

(b) **Requirements**: [R1] DisCoM shall manage Juridical Cases Database; [R2] The JC Requests Transactions (JCRT) shall be stored in JCRT-Table; [R3] DisCoM shall store the JC statuses allowing to provide a client with a JC status (JCS-Table); [R4] For each JC, only the latest update shall be stored in the JCS-Table; [R5] Upon receiving a JCR, the DisCoM shall notify all other parties regarding the update (for simplicity reasons we assume that updates non-relevant for a client shall be filtered out on the client side); [R6] Any sequence of JCRT-s regarding any sub-set of JC resources shall lead to a consistent state of the JCDB

This project objective is to implement **SiFra-DisCoM** (Simulation Framework for DisCoM) allowing to simulate and test various execution scenarios and insure that any sequence of JCRT-s regarding any sub-set of JC resources shall lead to a consistent state of the database. Students shall be requested to implement following: (a) **UI** for scenario creating, (b) **Format** for scenario storage, (c) **Scenario Player**, to execute scenarios and store their results to DB and logs, (d) **Validator**, for executed scenarios consistency checking.

The development shall be done in Java (optionally – Python), using JDBC.

**Programming Languages and Development Platform:** Java, Databases

**Courses:** 234122 - introduction to Systems Programming

**Recommended Background:** 236370 - Parallel and Distributed Programming, 236363 - Database Systems
Company: IBM  
Project Domain: Document classification  
Project: Document Classification for regulatory compliance  
Supervisor: Micha Moffie  

Project Description: Many companies today are facing the burden of regulations ranging from privacy regulations (e.g. the EU Data protection directive) to export regulations (e.g. US International Traffic in Arms Regulations). As the amount of data collected increases and the workforce is becoming more global, adhering to these regulations is becoming more and more complex. In this project we will look at ways to alleviate some of the regulatory burden and begin to automate some of the regulatory processes.  

Document classification techniques have long been applied in areas such as language identification and spam detection. In this project we will apply document classification techniques to determine whether a document should be subject to regulation. For example, does the document contain personal or sensitive data? Or, does the document contain munitions or plans for a guided missile?  

The scope of the project will be limited to several aspects of a single regulation. Students are expected to develop different heuristics that are able to address those aspects of the regulation (e.g. private information or technologies under export control) and identify if and where these exist in a document. Part of the project will include developing a prototype showcasing the effectiveness of the proposed heuristics.  

Programming Languages and Development Platform: Flexible  
Courses:  
- 234118 Computer Organization and Programming
- 234122 Introduction to Systems Programming
- 234218 Data Structures 1

Background in natural language processing – a plus.

# PROJECT 31

**Company:** IBM  
**Project Domain:** Adaptive bitrate video streaming, HLS, DASH  
**Project:** Live video crawler using the HLS and DASH adaptive bitrate video streaming standards  
**Supervisors:** Ophir Azulai and Udi Barzelay  

**Project Description:** Adaptive bitrate video streaming is a method of video streaming over HTTP where the source content is encoded at multiple bit rates, then each of the different bit rate streams are segmented into small multi-second parts. The streaming client is made aware of the available streams at differing bit rates, and segments of the streams by a manifest file. When starting, the client requests the segments from the lowest bit rate stream.

HLS and DASH are adaptive bitrate video streaming standards which are used today in the industry. The hls.js is a java script open source for HLS player and dash.js is a java script open source for DASH player.

FFmpeg is an open-source software project that produces libraries and programs for handling multimedia data.

In this project we will develop a live video crawler module, running in Node.JS that connects to a live video source using the hls.js / dash.js open source, receives the video segments, decode it using FFmpeg and detect text in the images using a text extraction service from IBM. The module will support multiple sessions in parallel.
Programming Languages and Development Platform: Java, Node.js, JavaScript
Courses: Software Engineering Methods (234321) or Object Oriented Programming (236703)

PROJECT 32

Company: IBM
Project Domain: Cloud Computing, Big Data
Project: Serverless Big Data Analytics
Supervisor: Alex Glikson
Project Description: hottest trends in Cloud Computing and Big Data – the "serverless" computing paradigm, and the interactive data science notebooks.

The serverless paradigm [1] emerged in the last 1-2 years, referring to an architecture where cloud-based applications don’t have traditional ‘backend’ components, but rather rely on 3rd-party micro-services as well as a collection of small, purpose-built ‘functions’ comprising the application’s business logic. These functions are hosted on a new generation of cloud-based frameworks, such as Amazon Lambda or IBM OpenWhisk, and are typically invoked in an event-driven manner. OpenWhisk [2, 3] is fully open source, developed by a vibrant and growing community of developers.

The interactive notebooks became the preferred way to interact with Big Data tools and technologies. A data scientist can use the intuitive graphical user interface of a notebook such as Jupyter [4] or Zeppelin [5] to develop and run short data processing and visualization programs, often interacting with powerful computational clusters at the backend (e.g., leveraging Apache Spark).
The goal of this project is to enable seamless offloading of computations from a notebook (such as Jupyter) to a serverless platform (such as OpenWhisk). This will require implementation of a new notebook ‘kernel’ [6], taking care of the interaction with the serverless backend along the life cycle of notebook applications, as well as various additional tweaks and enhancements across the different layers of the stack (in order to achieve smooth end-to-end integration and the desired user experience). Moreover, a performance analysis of the new architecture will be performed, using relevant Big Data benchmarks.


**Programming Languages and Development Platform:** Python, Java, JavaScript. Eclipse, Linux, Docker, cloud

**Recommended Background:** Excellent hands-on skills

---

**PROJECT 33**

**Company:** IBM

**Project Domain:** Cloud systems, security, machine learning

**Project:** Cognitive Cloud Shield

**Supervisors:** Shelly Garion and Yaron Weinsberg

**Project Description:** With the advance of machine learning algorithms and specifically deep learning, one can easily use computer programs to identify objects in pictures and documents.
People store all sorts of data in the cloud. One of the major concerns is that sensitive or private data is also stored, either by mistake or on purpose, thus may be leaked and used for evil purposes.

The project goal is to design and implement a system for intercepting data access requests made against some cloud based data store (such as S3, Swift, MongoDB) and block requests that may access sensitive data. We would like to either block a request, or blur the sensitive data in the image according to a simple policy based on who is the user invoking the request, whereas the data owner will always be able to get the original image. We would like to support various sensitive data types, such as: "faces", "passports", "receipts", "W2-tax form".

**Programming Languages and Development Platform:** The code will be developed for NodeJS/Express or Python Flask web toolkits.

We plan to use open source packages such as:


**Courses:** Introduction to Computer Science, Programming Languages, Introduction to Computer Networks, Introduction to Machine Learning (preferable)

---

**PROJECT 34**

**Company:** IBM

**Project Domain:** Cloud, PAAS, Docker

**Project:** Video surveillance service using FFmpeg and Docker on IBM Bluemix cloud

**Supervisors:** Ophir Azulai and Yevgeni Burshtein

**Project Description:** Bluemix is an implementation of IBM's Open Cloud Architecture based on Cloud Foundry, an open source Platform as a Service (PaaS). Bluemix delivers enterprise-level...
services that can easily integrate with your cloud applications without you needing to know how to install or configure them. Bluemix provides access to a wide variety of services that can be incorporated into an application.

Docker containers wrap up a piece of software in a complete filesystem that contains everything it needs to run: code, runtime, system tools, and system libraries – anything you can install on a server. This guarantees that it will always run the same, regardless of the environment it is running in.

FFmpeg is an open-source software project that produces libraries and programs for handling multimedia data.

In this project we will develop a basic video surveillance service on Bluemix using Docker and FFmpeg. In the client side we will have a gateway which will get the video stream from the camera, divide it to short segments and upload it to the cloud. The server side will receive the video and detect motion in it using an existing algorithm that IBM will supply. When a motion is detected, an event will be sent to a test web application.

The motion detection algorithm will run in a Docker container and video segments will be sent to it using the Kafka message queue.

Video decoding will be done using FFmpeg.

**Programming Languages and Development Platform:** Java, Golang, Linux

**Courses:** Software Engineering Methods (234321) or Object Oriented Programming (236703)
PROJECT 35

Company: Amdocs

Project Domain: Open Source, WebRTC, Unified communication, IoT

Project: Bridge the gap between different communication hubs

Supervisor: Yuval Lib

Project Description: The project aim is to connect widely used communication hubs like Hubot or Zapier (you can of course be creative and suggest to address a different hub)

In this project you will use matrix.org open standard and contribute to this cutting edge open source project by developing the plugins to a widely used leading communication hubs

Background: Matrix is an open standard for interoperable, decentralized, real-time communication over IP. It can be used to power Instant Messaging, VoIP/WebRTC signaling, Internet of Things communication - or anywhere you need a standard HTTP API for publishing and subscribing to data while tracking the conversation history.

Matrix defines the standard, and provides open source reference implementations of Matrix-compatible Servers, Clients, Client SDKs and Application Services to allow anyone to create new communication solutions or extend the capabilities and reach of existing ones.

More details on Matrix can be found here: [http://matrix.org/](http://matrix.org/)

Glimpse of the Matrix community sprit can be found in the Matrix blog: [http://matrix.org/blog/2015/12/25/the-matrix-holiday-special/](http://matrix.org/blog/2015/12/25/the-matrix-holiday-special/)

Goal

Matrix goal is to act as a generic HTTP messaging and data synchronization system for the whole web and mobile: allowing people, services and devices to easily communicate with each other, empowering users to own and control their data and select the services and vendors they want to use.
Your Mission: In order to achieve that Matrix must seamlessly connect to other communication hubs. Our goal is to have bridges/connections/integrations between matrix and as many communication hubs as possible.

You mission is to implement such a communication hubs bridge:

- Join the matrix community of topnotch open source programmers
- Design, develop and deploy a new bridge between Matrix and another management service, such as:
  - Zapier - automate tasks between web apps
  - Hubot - an automation tool that can sync with other chat services
  - Nagios – open-source application that monitors systems, networks and infrastructure
  - IFTTT - Create simple connections between the products you use every day (bridge the bridge)

(Of course you are free to be innovative and suggest another communication hub – it is up to you to convince your mentor or the matrix community that it is necessary)

We will mentor you and guide you throughout the process. Your success is our success and we will make sure you will excel (and enjoy the journey)!

Recommended Background: Web programming, web services

PROJECT 36

Company: Amdocs

Project Domain: Open Source, WebRTC, Unified communication, IoT

Project: Bridge the gap between different management tools

Supervisor: Yuval Lib
**Project Description:** The project aim is to connect widely used management tools like Trello or Basecamp (you can of course be creative and suggest to address a different management tool). In this project you will use matrix.org open standard and contribute to this cutting edge open source project by developing the plugins to a widely used leading management tools.

**Background:** Matrix is an open standard for interoperable, decentralized, real-time communication over IP. Matrix can be used to power Instant Messaging, VoIP/WebRTC signaling, Internet of Things communication - or anywhere you need a standard HTTP API for publishing and subscribing to data while tracking the conversation history.

Matrix defines the standard, and provides open source reference implementations of Matrix-compatible Servers, Clients, Client SDKs and Application Services to allow anyone to create new communication solutions or extend the capabilities and reach of existing ones. Matrix also has an active community of contributors to its open source projects that will be happy to guide and support you along the way.

More details on Matrix can be found here: [http://matrix.org/](http://matrix.org/)

Glimpse of the Matrix community sprit can be found in the Matrix blog: [http://matrix.org/blog/2015/12/25/the-matrix-holiday-special/](http://matrix.org/blog/2015/12/25/the-matrix-holiday-special/)

**Goal**

Matrix goal is to act as a generic HTTP messaging and data synchronization system for the whole web and mobile: allowing people, services and devices to easily communicate with each other, empowering users to own and control their data and select the services and vendors they want to use.

**Your Mission**

In order to achieve that Matrix must seamlessly connect to other management tools. Our goal is to have bridges/connections/integrations between matrix and as many management tools as possible.
You mission is to implement such a bridge:

- Join the matrix community of topnotch open source programmers
- Design, develop and deploy a new bridge between Matrix and another management service, such as:
  - Basecamp - web-based project-management tool
  - Trello - Trello keeps track of everything, from the big picture to the minute details

(Of course you are free to be innovative and suggest another management tool – it is up to you to convince your mentor or the Matrix community that it is necessary)

We will mentor you and guide you throughout the process. Your success is our success and we will make sure you will excel (and enjoy the journey)!

**Recommended Background:** Web programming, web services

---

**PROJECT 37**

**Company:** Amdocs

**Project Domain:** Open Source, WebRTC, Unified communication, IoT

**Project:** *Where do you store your images? Why should it matter! Bridge the gap between different media services*

**Supervisor:** Yuval Lib

**Project Description:** The project aim is to connect widely used media service like Google Images or Giphy (you can of course be creative and suggest another media service to address) so anyone can store and share images and movies seamlessly with anyone across any platform.

In this project you will use matrix.org open standard and contribute to this cutting edge open source project by developing the plug-ins to a widely used leading media services.
Background: Matrix is an open standard for interoperable, decentralized, real-time communication over IP. Matrix can be used to power Instant Messaging, VoIP/WebRTC signaling, Internet of Things communication - or anywhere you need a standard HTTP API for publishing and subscribing to data while tracking the conversation history.

Matrix defines the standard, and provides open source reference implementations of Matrix-compatible Servers, Clients, Client SDKs and Application Services to allow anyone to create new communication solutions or extend the capabilities and reach of existing ones. Matrix also has an active community of contributors to its open source projects that will be happy to guide and support you along the way.

More details on Matrix can be found here: http://matrix.org/

Glimpse of the Matrix community spirit can be found in the Matrix blog:
http://matrix.org/blog/2015/12/25/the-matrix-holiday-special/

Goal

Matrix goal is to act as a generic HTTP messaging and data synchronization system for the whole web and mobile: allowing people, services and devices to easily communicate with each other, empowering users to own and control their data and select the services and vendors they want to use.

Your Mission: In order to achieve that Matrix must seamlessly connect to external media services. Our goal is to have bridges/connections/integrations between matrix and as many services as possible.

Your mission is to implement such a bridge:

- Join the matrix community of topnotch open source programmers
- Design, develop and deploy a new bridge between Matrix and another media services, such as:
  - Giphy - search, share, and discover the world's best GIFs
Google Image Search - Comprehensive image search by Google.
(Of course you are free to be innovative and suggest another media service – it is up to you to convince the mentor or the Matrix community that it is necessary)
We will mentor you and guide you throughout the process. Your success is our success and we will make sure you will excel (and enjoy the journey)!

Recommended Background: Web programming, web services

---

**PROJECT 38**

**Company:** Amdocs

**Project Domain:** Open Source, WebRTC, Unified communication, IoT

**Project:** Connect Facebook Messenger and WhatsApp: Bridge the gap between different messaging services

**Supervisor:** Yuval Lib

**Project Description:** The project aim is to connect widely used messaging service like Facebook Messenger and WhatsApp (you can of course be creative and suggest another chat service to address), so each user can select the service she likes and still exchange message between one another.

In this project you will use matrix.org open standard and contribute to this cutting edge open source project by developing the plugins to a widely used leading messaging services.

**Background:** Matrix is an open standard for interoperable, decentralized, real-time communication over IP. It can be used to power Instant Messaging, VoIP/WebRTC signaling, Internet of Things communication - or anywhere you need a standard HTTP API for publishing and subscribing to data while tracking the conversation history.
Matrix defines the standard, and provides open source reference implementations of Matrix-compatible Servers, Clients, Client SDKs and Application Services to allow anyone to create new communication solutions or extend the capabilities and reach of existing ones.

More details on Matrix can be found here: http://matrix.org/

Glimpse of the Matrix community sprit can be found in the Matrix blog: http://matrix.org/blog/2015/12/25/the-matrix-holiday-special/

**Goal**

Matrix goal is to act as a generic HTTP messaging and data synchronization system for the whole web and mobile: allowing people, services and devices to easily communicate with each other, empowering users to own and control their data and select the services and vendors they want to use.

**Your Mission:** In order to achieve that Matrix must seamlessly connect to other messaging services. Our goal is to have bridges/connections/integrations between matrix and as many messaging systems as possible.

You mission is to implement such a bridge:

- Join the matrix community of topnotch open source programmers
- Design, develop and deploy a new bridge between Matrix and another messaging service, such as:
  - HipChat - XMPP based group chat and video chat built for teams.
  - Whatsapp - cross-platform mobile messaging app
  - Twitter - online social messaging supporting 140-character messages called "tweets"
  - Facebook Messenger – social texting app

(Of course you are free **be innovative** to suggest another messaging system – it is up to you to convince your mentor or the Matrix community that it is necessary)
We will mentor you and guide you throughout the process. Your success is our success and we will make sure you will excel (and enjoy the journey)!

**Recommended background:** Web programming, web services

### PROJECT 39

**Company:** Amdocs  
**Project Domain:** Big Data, analytic, diagnostic, REST  
**Project:** Big Data - Diagnostic package collector  
**Supervisors:** Idan Raphael and Menachem Kaplan

**Project Description:** Big data analytics (BDA) is an Amdocs framework for providing data centric applications. It is based mainly on Hadoop and its standards ecosystems tools.

The aim of the project is to provide a diagnostic package collector in order to create a utility that enable to collect diagnostic information from the Hadoop cluster and pack it in an archive, for two purposes:

1. Offline analysis of failures  
2. Analysis of performance bottlenecks

The tool should receive a timeframe to work on, and a diagnostic level (verbose / trace / error)

Then it should connect to the Hadoop cluster and other components (Elastic search cluster for example), and pull the diagnostic information from it using various protocols (mostly REST and java clients, but may be others, such as SSH / SFTP, etc.)

After pulling the data additional processing might be required (in error level, for example, error messages should be extracted from the logs)

Then the diagnostic information need to be stored in a hierarchical structure and packed in an archive
The tool should also be extendable with custom collectors, for future use, so, if for example data need to be collected from a new system, the user can inject new collector that the tool will use in order to collect data from the new system.

Implementation should be done with Java or Groovy

**Programming Languages and Development Platform:** Java or Groovy

**Technical Expertise required:** Java and / or Groovy

---

**PROJECT 40**

**Company:** Amdocs

**Project Domain:** Predictive analytics algorithms, Big Data

**Project:** From Information to Optimization by Predictive analytics algorithms

**Supervisor:** Yuval Balak

**Project Description:** From Information to Optimization by Predictive analytics algorithms

In this project we will define a set of rules to check and predict main Program pillars and optimize them – for better planning and ongoing execution, allowing early detection and mitigations for future issues.

“Analytics becomes a key competitive weapon- Leaders are addressing the diversity of data and unlocking the value of data via algorithms tuned to anticipate and deliver customer value” (Forrster)

**Background:** Today we are capturing in the Program management DWH various data segments from the different Amdocs system of the development life cycle from the pre sales through going live to production up to stabilizing the systems to customer satisfaction and support in customer premises.
• We would like to use the existing information systems to solve problems in the software development lifecycle
• Predictive analytics algorithms that can allow finding patterns in data that can make increasingly accurate predictions about the future of the business based on the relevant data so we will be able to (e.g.)
  o Early risks detection
  o Better Planning & budget optimization
    ▪ Resource optimization
    ▪ Pipeline optimization
    ▪ Activities predicted effort and final cost
    ▪ Project potential delays for contractual MSs
  o Execution
    ▪ Identify code areas with high risk for defects
    ▪ Defects closure models
  o And Other

Goal: The goal of this project is to be able to run an analytic tool that will be able to do data mining on the DWH and will be able to find prediction trends, behavior patterns and anomalies preventing future potential issues.

Recommended background:
• Knowledge in data structures, and SQL
• Predictive modeling
• Machine learning
• Data Mining
PROJECT 41 – CANCELED!

PROJECT 42

Company: Amdocs

Project Domain: Integration hub with multiple chatbots providers

Project: Chatbots Hub

Supervisors: Sébastien Bariou and Yinon Amir

Project Description:
- Identify the various chatbots protocols existing today.
- Design a solution that will
  - Easily on board chatbots systems
  - Assess which one should be requested (optional)
  - Present the answers in a unified way through a web interface

Solution components:
- Chatbots bus
- Chatbots connectors / configurator
- Requests router (optional)
- Web integration SDK
- Simple web front end

Recommend Background: Web applications development.
PROJECT 43

**Company:** Amdocs

**Project Domain:** Web application development mechanism

**Project:** Text Provider for web applications

**Supervisor:** Alaa Bishara

**Background:**
- For a Single Page web Application / widget, Textual Resources are available as key - value pair in a JSON / XML file in app source control or they are hardcoded in the HTML DOM.
- Any change on textual property, will require IT / Development involvement:
  - File Check-out, edit, check-in and run build / apply HF on the environment.
- Customers need to provide the applications texts upfront.
- Customers not always ready to provide and asking for changing text during development.
- No Preview / Immediate feedback, causing out of bounds texts / rollback change.
- Multilingual sites are more complex, usually handled by external parties, typically excel export / import to the application source control and proceed with deployment process to preview changes.
- Process is out of customer control and not flexible.

**Project Description:** Create web application with friendly UI, allows user to do add / update the web application texts by configuration, without compiling and delivery.

Application generate a minified is [TBD] library uses JSON object contains the set of key/value pairs to be included in the web page / website template.

Developers will be able to get the values of the key’s using available function in the included library.

**Solution components:**
- Small website application developed by
  - Client - AngularJS/ReactJS
  - Server – NodeJS/Java

**Recommend Background:** JavaScript development, JSON/HTML/CSS knowledge.

---

**PROJECT 44**

**Company:** Amdocs

**Project Domain:** Web and Mobile UI, JavaScript Application.

**Project:** *Fluid Page Layout without WCMS – Dynamic UI presentation*

**Supervisor:** Ron Weisgross

**Background:** WCMS (web content management system) there is software tool that provide capability to manage web site content and presentation layout layer with very little knowledge of web programming.

Currently most of web site doesn’t have WCMS.

Following reference to few populated WCMS tools:

- [https://dotcms.com/](https://dotcms.com/)
- [https://www.drupal.org/](https://www.drupal.org/)

**Project Description:** Provide framework this will support web site layout changes as remove, add, moving web components/widget without using a WCMS

To develop a mechanism to layout a page dynamically which defining the template structure (header, footer, 2 columns 30% 70%, 3 default components etc...) the page will contain a list of components available to be used and a default layout and the user will be able to dynamically remove components and add components from the available list and to change the order by drag and drop
Solution components:
- Small website application developed by
  - Client - AngularJS/ReactJS
  - Server – NodeJS/Java

**Recommend Background:** JavaScript development, JSON/HTML/CSS knowledge

---

**PROJECT 45**

**Company:** Amdocs  
**Project Domain:** limited time promotions using Snapchat  
**Project:** Snapchat advertisement platform  
**Supervisor:** Ron Weisgross

**Background:** Snapchat is very popular social platform which allows media to be exchanged between parties with expiration time, it can be used for marketing –
- Offer rewards and coupons
- Promote new products
- Inform about special sale events
- Etc...

**Project Description:** Develop a digital platform which integrates with Snapchat and sends media files to a selected audience
- Design a UI and business flows
- Implement integration framework with Snapchak
- Show dashboard of the advertisements and exposure rates

Solution components:
- Implement Snapchat integration
- UX design and development
- Full stack web framework

**Recommend Background:** Web applications development, system integration, databases

---

**PROJECT 46**

**Company:** Amdocs

**Project Domain:** Web\Mobile & Analytics application

**Project:** SMB Marketplace Employee Cockpit

**Supervisor:** Shachar Prizat

**Project Description:** This exciting project is about developing an advanced Small Medium Business (SMB) platform that turn the current client from limited app launcher to central business employee cockpit tool that will enable mobile application launch, wider options of Telco specific data presentation and advanced business management features. The tool will improve the business employee productivity and the collaboration inside the organization.

As part of the project the students will develop the client framework + few sample applications adopting it.

**General explanation:** The current Amdocs SMB Marketplace offering is an end to end platform & eco system that connects Telecommunication Service Providers with small and medium business (SMB) and allows the latter to consume all their IT apps and services in a cloud SaaS model and serve as a app launcher.

The platform includes web & mobile client application that is using Microsoft leading IDaaS (Identity as a Service) solution – AAD, in order to allow secured access to the organization applications via mobile and web apps.
Currently we are using azure AD access panel (http://myapps.microsoft.com) as the SMB employee app launcher.

Background details:

- You can learn about Microsoft IDaaS solution in here: https://azure.microsoft.com/en-us/services/active-directory/

- You can learn about Microsoft access client in here: https://azure.microsoft.com/en-us/documentation/articles/active-directory-saas-access-panel-introduction/

Here is an illustration of the cockpit – one possible implementation:

**Features:**

- Web app + mobile app (android + ios)
- CSP Branded Portal Sign on using AAD user
- The portal: Single page with few areas or different pages with tabs:
  1. Business App launcher: Get list of authorized apps from AAD and display them in a tiles display. Show last login datetime on the tile and who else is logged in (from my SMB)
  2. My business data: my main business status (based on role) taken from few apps – dynamically built per user based on the SMB apps list and role – not personalized. Number of orders today, my next delivery (if I am a field employee), team calendar

Can be based on Cortana analytics + sharepoint widgets or deeplinking technology – We are open to innovative suggestions

1. Personal page: Add launch to personal apps, notes ...
2. Promotions from the CSP
3. Ads using 3’rd party ad provider – pending on the price plan the SMB purchased.
• The portal pages will be built dynamically based on the list of apps purchased (and apps widgets will be resized accordingly)
• Each widget will also allow navigation to the app itself (can be to the home page of the app or to inner page if possible using deeplink technology)
• App launcher desktop – federation based SSO using AAD setup (graph api).
• App launcher mobile – if public app available (app metadata) then check if installed over the phone.
  If yes – launch it (SSO) – I guess using deeplinking technology
  • If not – prompt for installation + link to the public/local MDM store to install it and launch it post installation. User can choose to skip and then launch the web app

Recommended background: Web Development, Mobile Development, Azure cloud development, deeplinking, SSO, Data analytics, Microsoft Cortana, Rest APIs

PROJECT 47

Company: Amdocs
Project Domain: Cognitive analytics, machine learning, automation, AI, Robotic Process Automation
Project: Cognitive Automation for Productivity (CAP system)
Supervisor: Yariv Hammer
Project Description:
Cognitive Automation for productivity: In this project we will build a system of desktop clients, each monitoring all activity done by a user on a machine, and through machine learning techniques produce an automated report of the tasks which are candidates for automation
through bots. The system will aggregate the data collected from a group of clients to provide prioritization of the bots based on the overall productivity benefits.

**Background:** Artificial Intelligence and machine learning are being adopted in enterprise environments, such as telecoms and financial services, in a wide array of business and operational domains. In this project we will focus on automation that drives productivity.

Robotic Process Automation (RPA) is a growing domain of which deals with the application of technology that allows employees in a company to configure computer software or a “robot” to capture and interpret existing applications for processing a transaction, manipulating data, triggering responses and communicating with other digital systems.

We would like to expand the notion from using bots to cognitive AIs that creates bots. The focus of the project is not the platform to record and deploy the bots, but on the cognitive layer that defines when bots should be built and what they should do.

**CAP system functionality:**

- **Phase 1** - build a desktop client that monitors and record all user activity on a machine and measuring the time of each activity
- **Phase 2** – create the cognitive layer for understanding the data, identifying repeatable tasks which are easily automated, including the ability to elaborate the flow of activities done in the task
- **Phase 3** – produce a report for each client on a periodic bases (daily, etc) of all the tasks which are candidates for automation
- **Phase 4** – aggregate the data of multiple clients to identify tasks that are common across the organization and prioritize bot creation according to productivity gains

**Goal:** The goal of this project is to be able to run desktop clients that identify tasks which are candidates of automation, including detailed description of the flow of each task, and produce a report for the organization of the top tasks that bring the most productivity benefits.
The output of the project will be the input of the robotic process automation platform that actually builds and runs the bots (out of scope for this project)

**Recommended background:** Machine learning, Client side monitoring, Analytics, Process modelling

---

**PROJECT 48**

**Company:** LogicBlox

**Project Domain:** Machine Learning, Databases

**Project:** *A Declarative System for Engineering Deep Learning*

**Supervisor:** Dr. Nikolaos Vasiloglou

**Project Description:** Deep Learning (DL) is among the topics that arouse the most excitement in the computer-science community nowadays. The major benefit of DL is the ability to learn complex functions from a highly rich space of computations (neural networks). Consequently, DL is able to operate over basic, low-level details of information (raw features), and allows to avoid the tedious (and domain-specific) challenge of feature engineering. DL has been so far successful in a handful of domains, such as image recognition and natural language processing, where labeled data is abundant due to modern social and technological trends. Moreover, modern hardware design, such as the GPU family, allow to conduct the heavy computation that DL training entails.

Engineering a task-specific solution using DL technologies requires expertise in neural networks, mastery of specialized libraries for linear algebra such as Signa and TensorFlow, and the transformation of data from the semantic database (e.g., people, reviews, words) into the algebraic representations of the libraries. The goal of this project is to drastically reduce the
amount of engineering and the level of expertise required for building DL applications. In particular, in the project you will build the language and translation component for constructing neural networks over a logic-based database, namely LogicBlox (www.logicblox.com). As a result, developers will be presented with a uniform language for representing the database, query the data for raw features, building the network, and using the trained model for prediction. We will focus on some specific forms of DL (e.g., feed-forward networks for multi-classification), and some specific tasks from actual client engagements of LogicBlox.

**Programming Languages and Development Platform:**
- c++ / python
- LogicBlox

**Required background:**
- Machine Learning
- Databases or Logic

---

**PROJECT 49**

**Company:** LogicBlox

**Project Domain:** Machine Learning, Data Mining

**Project:** A Tool for Data Quality Management in Machine Learning

**Supervisor:** Dr. Nikolaos Vasiloglou

**Project Description:** The abundance of Big Data resources opens up unprecedented opportunities for Machine Learning (ML) applications, towards a variety of tasks such as automation of processes (Artificial Intelligence) and predictive analysis (e.g., demand forecasting
and disease prediction). However, such data comes with significant challenges to ML libraries. One of the greatest challenges is that of controlling the quality of data. Low data quality has several basic reasons in the context of Big Data. Repositories are often collected from sources with limited content control, and they undergo nontrivial automated processing before being served to the ML library at hand. For example, data may be integrated from resources that disagree on formats, or contain overlapping information. Data may be generated via imprecise methods such as signal analysis, image recognition, and natural language processing. And data may undergo transformations that fail on certain values or character encodings.

While low data quality is an old problem, the high volumes processed nowadays often make it impossible to detect and correct errors. In fact, one typically observes the errors only when the ML library mal-behaves. The easy case is when exceptions are thrown. The hard case is when the learned prediction model behaves in an unintuitive manner (e.g., reducing the price of a product leads to reduction of sales). In this project, you will develop a tool for dramatically facilitating the control of data quality in the context of machine learning applications. In particular, the tool will focus on the management of large volumes of training data, and will serve two main functionalities.

1. The tool will gather and visualize statistics on columns, such as histograms, curve fitting, outlier detection, correlation among columns (e.g., one column is generally monotonic in the other), and integrity rule mining.
2. The tool will allow the ML developer to phrase expected behavior patterns (e.g., sales go up when price goes down), and consequently, to detect unexpected behavior in the resulting ML model, and to evaluate the compatibility of different ML models with the developer’s expected behavior.
The project will be co-supervised by a professor at Technion CS (Prof. Benny Kimelfeld), and a researcher at LogicBlox, Atlanta (Dr. Nikolaos Vasiloglou). The tool will be developed within one or more specific deployments of active customers of LogicBlox.

**Programming Languages and Development Platform:**
- c++ / python
- Web UI design

**Useful background/courses (not a formal requirement):** Machine Learning, Statistics, Databases, Logic, Artificial Intelligence, Web Design.